Non Intrusive Measures for Determining the Minimum Field of View for User Search Task in 3D Virtual Environments

Zahen Malla Osman, Jérôme Dupire, Alexandre Topol, Pierre Cubaud
Centre d'Etude et de Recherche en Informatique et Communications
Conservatoire National des Arts et Métiers
Paris, France
Email: {zahen.malla_osman, dupire, topol, cubaud}@cnam.fr

***Abstract*—In this paper we report on an experiment to determine the minimum field of view (FOV) that permits the user to perform an effective search task in a 3D virtual environment (VE), by analyzing how the user controls the virtual camera. Our study exploits a model based on the use of several novel non intrusive temporal and quantitative measures of visual attention such as fixation, gaze and movement. Seven out of ten measures gave significant results with the same findings.**

Keywords-Field of view; virtual environment; video games; visual attention.

I. Introduction

Visual attention is the ability of a vision system, whether human or artificial, to quickly select the most pertinent information from the environment in which it operates [1].

Eye tracking has been used to measure visual attention for many years. It is the process of using sensors to localize the position and the behavior of the eyes. It helps us to determine what a person is looking at, what she is not looking at, but also what she does and does not pay attention. Through eye tracking systems we can provide many visual attention measures such as fixation, gaze and movement in order to analyze users' ocular behavior.

A principal means of interacting with 3D VEs, in the case of video games, for example, is the use of the virtual camera, which is relatively easy to access and manipulate via game engines. The use of this virtual camera can show interesting results for non-invasive study and characterization of user behavior - especially in the absence of eye tracking systems, which can sometimes be unavailable.

Our work is focused on the FOV effect of the virtual camera for determining the minimum FOV that allows the user to perform an effective search task in a 3D VE.

In this contribution first we present related work. Next, we describe our experiment that analyzes user behavior in a 3D VE via the virtual camera. Finally, we summarize our paper and provide an outlook for future work.

II. Related Work

Gaming is an increasingly prevalent cultural pastime [2], and today the video game is one of the most popular types of software applications in the world. More than half of all Americans play video games, for example [3][4]. Video games can provide a framework for testing many types of attention measures [5], e.g. playing video games such as Pac Man can improve the reaction times of older adults [6].

During everyday interactions our eyes reflect a lot of information that affects our emotional and mental states. Eye movement data reflects moment-to-moment cognitive processes during task execution [7]. When we look at an object in space (e.g. a wall with windows and doors), our eyes concentrate much more on some parts of this object (e.g. one of the windows), while the other parts of the object may receive a little attention [8].

Studying ocular behavior in the context of human computer interaction (e.g. web browsing or video games [9][10][11][12]), allows us to identify and provide many indicators that can be used to evaluate user attention in order to improve the design of an user interface such as, for example, a digital library [13].

Much research have been conducted towards the study of ocular behavior during playing a video games. During a First Person Shooter (FPS) game, for example, the player pays more attention to the center of the screen around the reticule because he shoots enemies through the reticule; by contrast, the attention area is larger in an adventure game because the player is not constrained by any specific area of the screen [14][15].

There are many types of eye behaviors: fixation, being the moment when the eyes are relatively stationary, taking in or encoding information with a minimum duration of 100 milliseconds [16]; saccade, being the eye movement occurring between fixations with durations of approximately 150–200 milliseconds [17]; and gaze, being the moment when the eyes look at a display element [18]. When we look at an object in a visual display, we may make many fixations on this object. The number of these fixations sometimes represents the importance of the display area, but a large number of fixations represents a poorly designed interface [18].

To study user behavior in a 3D VE, we offer a search task to the user in order to know how this user interacts with the 3D VE: e.g. a task in which users have to find objects that have specified numbers displayed on them [11], or a task where the user must find a maximum number of hidden keys distributed in a 3D VE [19].

Our idea was to use the virtual camera of a 3D VE to examine several visual attention measures such as fixation, gaze and movement. The use of the 3D VE's virtual camera provides an indirect method for analyzing the FOV effect on user behavior, given that the useful FOV is the total area of the visual field within which individuals can obtain useful information without moving their heads or eyes [20][21].

In order to study and characterize user behavior in a 3D VE through the virtual camera, we selected several visual attention measures employed by Gibbs et al. [9]. The measures selected are expressed by the number of fixations, fixation duration and gaze duration. We also added a certain number of measures to give more information about how the user executes an effective search task in our 3D VE. The measures that we added are expressed by: the number of gazes, number of movements, movement duration, the sum total duration of all fixations per task, the sum total duration of all gazes per task, the sum total of all movements per task, and the total duration of each task.

Our goal is to examine the FOV effect of the virtual camera on user behavior and to determine the minimum FOV that allows the user to perform an effective search task in a 3D VE. FOV size is very important for rapid extraction and identification of information in the 3D VE. We consider an effective search task, in the context of our experiment, to consist of a simple navigation within the 3D VE for the purpose of finding all objects (e.g. hidden buttons distributed around the VE) using the least possible number of fixations and the shortest fixation duration; the least possible number of gazes and the shortest gaze duration; the least possible number of movements and the shortest movement duration; the shortest sum total duration of all fixations per task; the shortest sum total duration of all gazes per task, the shortest sum total of all movements per task; and the shortest total duration of each task. Our results provide information that can be of benefit to game designers, allowing them to improve gameplay, manage the difficulty of game environments and optimize the distribution of visual resources.

III. OUR EXPERIMENT

Gibbs et al. used an eye tracking system to determine whether the ocular behavior differs between newspaper websites and TV-oriented websites. They used several visual attention measures to test ocular behavior such as number of fixations, fixation duration and gaze duration. Within the contest of FPS video games, our research uses these measures employed by Gibbs et al. as well as our own measures to analyze user behavior, using the virtual camera of our 3D VE instead an eye tracker. The aim of our experiment is to test FOV effects on user behavior and to generate information for game designers to help them manage and adapt difficulty of a 3D VE according to user behavior.

The users in our experiment use a mouse and a keyboard to manipulate the virtual camera of our 3D VE like they would in a FPS video game (e.g. Half Life, Counter Strike). The measures employed in our experiment consists of various types such as: fixation, being a short pause in movement, represented quantitatively by the number of fixations (NF) and temporally by the fixation duration (FD), which varying between 100 and 300 milliseconds; gaze, which is the time spent looking at a display object, represented by the number of gaze (NG) and the gaze duration (GD),which starting from 300 milliseconds; and movement between two fixations or gazes, represented by the number of movement (NM) and the movement duration (MD), which starting from 100 milliseconds.

We also added four measures to those specified above: the sum total duration of all fixations per task (STDF), the sum total duration of all gazes per task (STDG), the sum total duration of all movements per task (STDM), and the total duration of each task spent by the user to achieve the required task (TD).

A total of 14 volunteers (10 male and 4 female) participated in this experiment. Their ages varied between 25 and 42 years, with a mean of 30. All participants are right-handed and healthy. The experiment was performed on a desktop personal computer with an LCD display with a resolution of 1920×1040 pixels.

A. *Procedure*

The purpose of the following experiment is to compute visual attention measures and to study the FOV effects on user behavior during a visual search task in a 3D VE, using the VE's virtual camera. Fig. 1 shows our 3D VE, which is a virtual art gallery similar to the static environment of Lee et al. [11]. We used Unity3D version 3.5 to create our 3D VE, including all the objects and the buttons. The virtual camera is positioned at the level of the eyes of user's avatar.

The participants were first invited to complete a short form to provide information including their name, age, gender and whether or not they often play FPS video games. Secondly, the participants were asked to perform a free navigation in the 3D VE with a FOV of 80°, simply navigating in the 3D VE and observing the virtual objects using the mouse and the keyboard to control navigation motion. This step was created as a training phase to teach manipulation of the virtual camera. The participants used the mouse to change the orientation of the virtual camera (yaw and pitch angles) and the keyboard to move the virtual camera. We used an 'AZERTY' format keyboard with the following key mapping: Z: forward, S: back, D: right, Q: left.

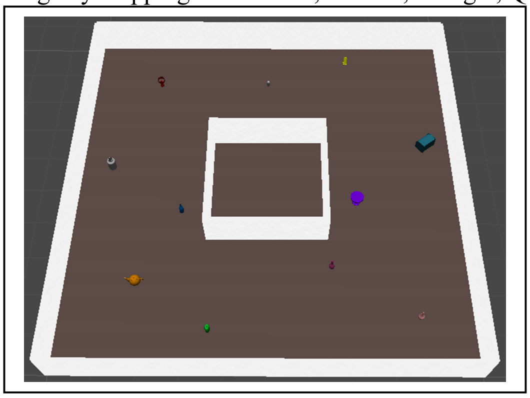

Figure 1. Our 3D virtual environment (the art gallery).

Finally, all participants were asked to perform a visual search task to find and validate hidden buttons in the 3D VE. They had to find ten buttons randomly distributed on the surfaces of objects of the 3D VE.

Each object in our VE contains one hidden button. The participant had to find these hidden buttons using the reticule area (a rectangle 250×150 pixel situated in the center of the center), and validate them by pressing the Space key. A number is displayed at the top left of the screen to indicate how many hidden buttons are left.

The participant was asked to repeat the search task six times, knowing that we had changed all the objects and the buttons positions and the FOV size of the virtual camera before each of the six attempts at the task (10° for the first attempt, 20° for the second, 30° for the third, 50° for the fourth, 80° for the fifth and 110° for the sixth attempt). The order of the attempts was randomly given to the participants in order to eliminate the adaptation effect. The purpose of changing the FOV size, from 10° to 110°, was to discover how the FOV affects user behavior and to determine the minimum FOV that enables the user to perform an effective search task in a 3D VE such as an FPS games, given that the default FOV in an FPS game ranges from 75° to110°.

B. Results

A one-way ANOVA was conducted to see whether the FOV of the virtual camera affected user behavior during the search task in our 3D VE. A total of 14 subjects took part of the experiment. We sought to discover whether there is a significant difference between the measures that we obtained by changing the FOV between 10°, 20°, 30°, 50°, 80° and 110°. We expressed our measures by way of a natural logarithm and tested the measures' normality using the Shapiro Wilk test. Then, we used the ANOVA test to analyze the variance between all our measures.

Table I shows the means, standard deviations and analyses of variance of all our measures. Our ANOVA results show a significant difference between certain measures used in our experiments when we changed the FOV; such as the NF, NG, NM, STDF, STDG, STDM and TD. However the FD, GD and MD don't show any significant difference.

To determine the minimum FOV that allows the user to conduct an effective search task within our 3D VE, we performed another ANOVA that examined all our measures between each FOV pair that we used in our experiment. The results of this ANOVA do not show a significant difference between FOV of 10° and FOV of 20° when using all the measures (p=0.0585), but they do show a significant difference between FOV of 10° and FOV of 30° (p=0.015*), 50° (p=0.012*), 80° (p=0.011*) and 110° (p=0.0002***). The ANOVA results also show a significant difference between FOV of 20° and the FOV of 110° (p=0.005**), but they do not show a significant difference between FOV of 20° and the other FOVs. Finally, this ANOVA shows that there is no significant difference between FOV of 30°, 50°, 80° and 110°.

Fig. 2 shows TD measure allocation by the FOV. The box plot presents the TD means of all participants in the six sizes of FOV. We observed that this measure decreases when the FOV increases. We found there to be not much change in user behavior when we used an FOV of 30°, 50°, 80° or 110°; however, an FOV of 10° or 20° shows a lot of change in user behavior. For example, users took a long time to accomplish the task when they used an FOV of 10° or 20°, while they took less time when they used a FOV of 30°, 50°, 80° or 110°.

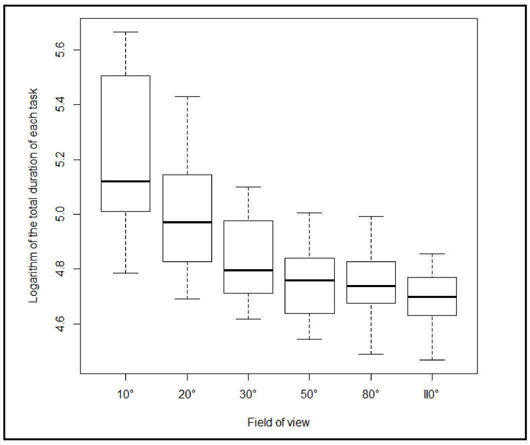

Figure 2. The total duration of each task (TD) by the field of view (FOV)

TABLE I. MEAN, STANDARD DEVIATIONS AND ANALYSES OF VARIANCE OF THE VISUAL ATTENTION MEASURES IN THE SIX SIZES OF FOV

	10°	**20°**	**30°**	**50°**	**80°**	**110°**	***F***	***p***
NF	2.15 (0.32)	1.84 (0.23)	1.78 (0.16)	1.66 (0.18)	1.60 (0.17)	1.57 (0.14)	14.92	<0.0001 ***
NG	1.93 (0.34)	1.66 (0.27)	1.50 (0.24)	1.45 (0.23)	1.43 (0.23)	1.40 (0.17)	8.99	<0.0001 ***
NM	2.35 (0.32)	2.06 (0.24)	1.96 (0.18)	1.87 (0.19)	1.82 (0.19)	1.79 (0.13)	12.68	<0.0001 ***
FD	2.23 (0.02)	2.23 (0.02)	2.21 (0.02)	2.22 (0.04)	2.22 (0.03)	2.21 (0.03)	1.12	0.358
GD	2.90 (0.09)	2.95 (0.12)	2.93 (0.11)	2.89 (0.10)	2.88 (0.11)	2.90 (0.12)	0.94	0.489
MD	2.44 (0.19)	2.53 (0.25)	2.49 (0.24)	2.51 (0.26)	2.57 (0.26)	2.47 (0.30)	0.41	0.838
STDF	4.37 (0.33)	*4.07 (0.21)*	3.99 (0.17)	3.89 (0.18)	3.82 (0.15)	3.78 (0.13)	15.81	<0.0001 ***
STDG	4.83 (0.40)	4.61 (0.33)	4.43 (0.32)	4.33 (0.31)	4.31 (0.32)	4.30 (0.27)	5.80	<0.0001 ***
STDM	4.79 (0.22)	4.59 (0.25)	4.45 (0.15)	4.37 (0.13)	4.39 (0.14)	4.26 (0.22)	13.91	<0.0001 ***
TD	5.20 (0.30)	4.99 (0.21)	4.85 (0.16)	4.76 (0.13)	4.75 (0.14)	4.69 (0.11)	14.84	<0.0001 ***

*** p <0.0001, ** p <0.001, * p <0.01, NF: the number of fixations, NG: the number of gazes, NM: the number of movements, FD: the fixation duration, GD: the gaze duration, MD: the movement duration, STDF: the sum total duration of all fixations per task, STDG: the sum total duration of all gaze per task, STDM: the sum total duration of all movement per task and TS: the duration of each test.

We also observed that the NF measure decreases when the FOV increases (see Fig. 3).

Fig. 4 shows the STDF measure according to the FOV. The STDF becomes convergent from an FOV of 30°.

The NG in the fourth task (FOV size = 50°) was high compared with the third, fifth and sixth tasks (respectively: FOV size = 30°, 80°, 110°) reflecting the impact of the hidden buttons positions (see Fig. 5).

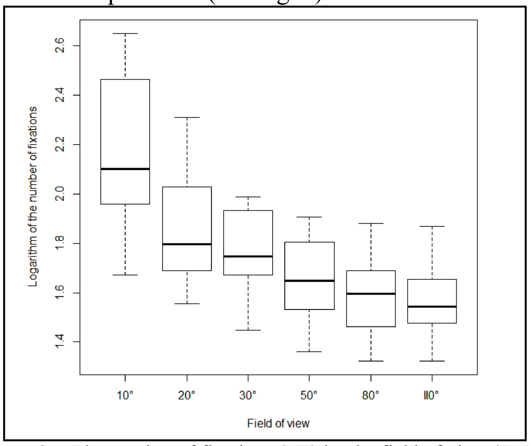

Figure 3. The number of fixations (NF) by the field of view (FOV).

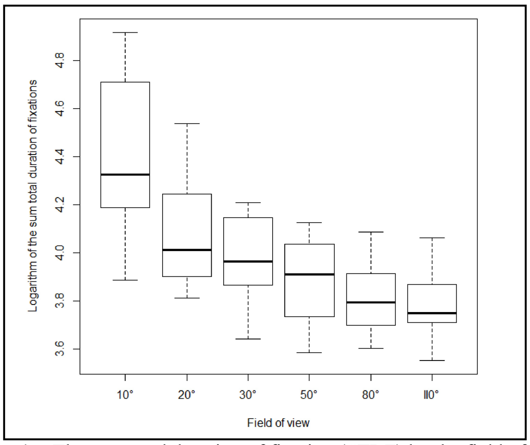

Figure 4. The sum total duration of fixation (STDF) by the field of view (FOV).

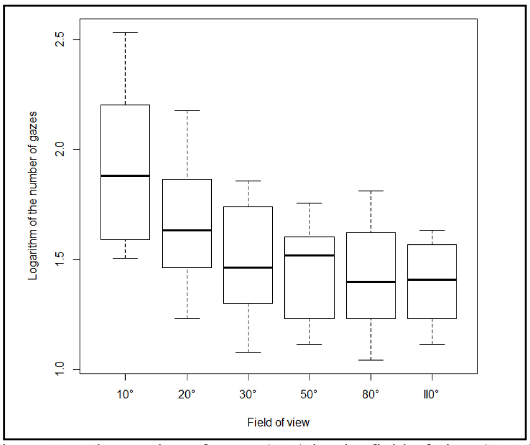

Figure 5. The number of gazes (NG) by the field of view (FOV).

Fig. 6 represents the STDG measure by the FOV. We observed through this measure that the STDG in the sixth task (FOV size = 110°) was high compared to the fifth task (FOV size = 80°), due to the use of a large FOV.

Fig. 7 shows the NM measure by the FOV. We also observed that the NM in the fourth task (FOV size = 80°) was high compared with the third, fifth and the sixth tasks (respectively: FOV size = 30°, 80°, 110°).

Fig. 8 shows the STDM measure by the FOV. We observed that there is a user in the third task (FOV size =30°) that is out the box plot. This is because the user had difficulty in manipulating the virtual camera.

After analyzing all our participants without taking into consideration their video game experience, we categorized our subjects into two groups: video game players (VGP) and non-video game players (NVGP). Fig.9 shows a comparison between the VGPs and the NVGPs using the TD measure. We found that the NVGPs took more time than the VGPs to achieve the required task.

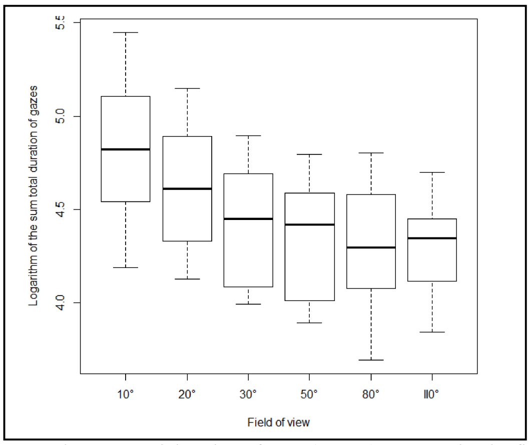

Figure 6. The sum total duration of gaze (STDG) measure by the field of view (FOV).

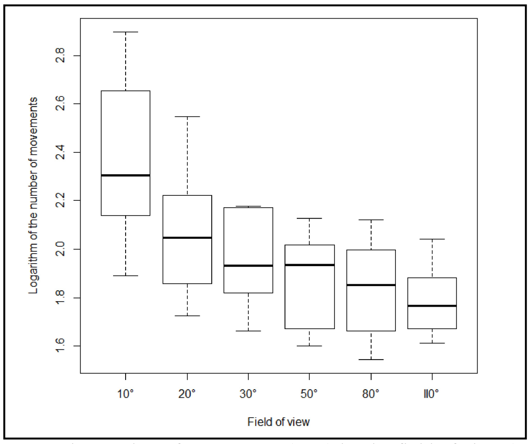

Figure 7. The number of movements (NM) by the field of view (FOV).

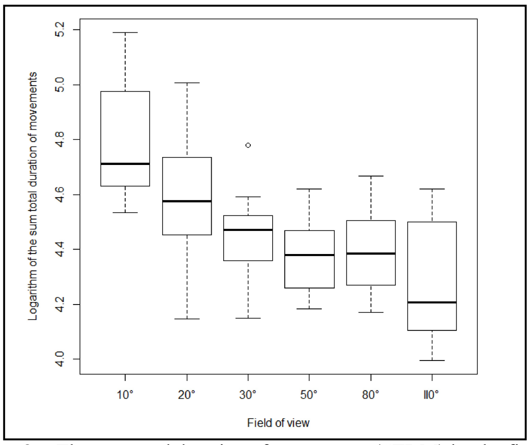

Figure 8. The sum total duration of movements (STDM) by the field of view (FOV).

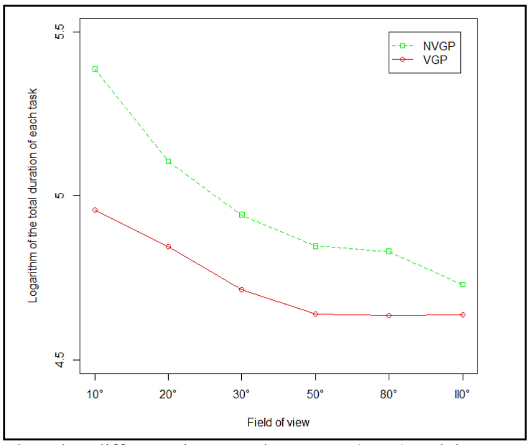

Figure 9. The sdifference between the gamers (VGP) and the non gamers (NVGP) using the total duration of each task (TD).

C. Discussion

We notice that in Table I there is a significant difference between the results based on most of our measures: the NF, NG, NM, STDF, STDG, STDM and TD. This difference between these measures is due to the change in the FOV (i.e. between 10°, 20°, 30°, 50°, 80° and 110°), where we observe that the FOV affects user behavior during navigation within a 3D VE. We note that these measures decrease as the FOV increases, e.g. the NF mean value for all the subjects had a natural logarithm of 2.15 when we used an FOV of 10°, and this NF decreased to 1.57 when we used an FOV of 110°. Additionally, the NG mean value for all the subjects had a natural logarithm of 1.93 when we used an FOV of 10°, and this NG decreased to 1.40 when we used an FOV of 110°. We observe also that the NM mean value for all our subjects had a natural logarithm of 2.35 when we used an FOV of 10°, and this NM decreased to 1.79 when we used an FOV of 110°. We found also that the STDF decreased as the FOV increased, where the STDF mean value for all the subjects had a natural logarithm of 4.37, and this STDF decreased to 3.78 when we used an FOV of 110°. The STDG mean value for all the subjects had a natural logarithm of 4.83, and this STDG decreased to 4.30 when we used an FOV of 110°.The STDM mean value for all the subjects had a natural logarithm of 4.79, and this STDM decreased to 4.26 when we used an FOV of 110°. Finally, the TD mean value for all the subjects had a natural logarithm of 5.20, and this TD decreased to 4.69 when we used an FOV of 110°. The decrease of measures is important for determining the FOV within which one can navigate effectively within a 3D VE.

The ANOVA preformed on each pair of FOVs allows us to define two groups of FOVs according to measure values: Group 1, with FOVs of 10° and 20°, and Group 2 with FOVs of 30°, 50°, 80° and 110°, given that there is not a significant difference between FOV 10° and FOV 20°, and between FOV 30°, 50°, 80° and 110°; but that there is a significant difference between FOV of 10° and FOV of 30°, 50°, 80° and 110° and between FOV of 20° and FOV of 110°. User behavior in Group 1 was less effective than user behavior in Group 2 because users in the Group 2 performed the search task quicker than users in Group1 with the least possible number of fixations; the least possible number of gazes the least possible number of movements; the shortest sum total duration of all fixations per task; the shortest sum total duration of all gazes per task, the shortest sum total of all movements per task; and the shortest total duration of each task. We observe that the user can use an FOV of 30° as a minimum FOV for performing the search task in a short time with minimum movement of the virtual camera. We observe also that the user can perform effective search task using this FOV of 30° in cases where we did not find much change in user behavior based on the virtual camera when she uses a large FOV such as 80° or 110°.

Finally, we see also in Fig. 9 that the NVGPs have spent more time than the VGPs to achieve a visual search task in a 3D VE, and therefore we can deduce that the VGPs perform better on the required task than the NVGPs because the VGPs are accustomed to playing video games.

IV. CONCLUSION

In this paper we have presented our model for the study and characterization of user behavior using the virtual camera of a 3D virtual environment, which is accessible in all game engines.

We used several visual attention measures to monitor user behavior. Our results, which are based on the use of a virtual camera of a 3D virtual environment, show differences in seven of ten measures of user behavior resulting from the field of view size.

The participants in our experiment could perform this effective search better when the visual attention measures values were smaller.

We have shown that the field of view of the virtual camera affects user behavior during navigation within a 3D virtual environment to complete a visual search task. Our quantitative and temporal measures were evaluated by changing the field of view size of the virtual camera in the virtual environment. We found that the user needed less time to achieve her visual search task if she used a large field of view. We also showed that the minimum field of view for performing an effective search task in a 3D virtual environment is 30°. Finally, we showed that the video game players perform better in the 3D.

Our model can be used in the context of video games to give additional information to game designers about the improvement of gameplay and management of difficulty by modifying the field of view of the virtual camera relative to the difficulty level or the player's needs.

For future work, we plan to prototype an First Person Shooter video game to show how our model can be of benefit to game designers. We will also test how our model can be used in the service of cognitive rehabilitation, specifically, for facilitating search tasks within and adapting the difficulty of a 3D virtual environment.

ACKNOWLEDGMENT

We would like to thank K. Neil, N. N. Keita and M. Ouattara, and all the people who accepted to test and comment our experiment.

REFERENCES

[1] H. Hugli, T.Jost, N. Ouerhani, "Model Performance for Visual Attention in Real 3D Color Scenes," Proc. of the first international work-conference on the Artificial Intelligence and Knowledge Engineering Applications: A Bioinspired Approach (IWINAC 05), 2005 ,vol. Part II, pp. 469–478, doi:10.1007/11499305_48.

[2] C. S. Green, D. Bavelier, "Action video game modifies visual selective attention," Nature Publishing Group, 2003, vol. 423, pp.534–537.

[3] A. Lenhart, S. Jones, A. Rankin Macgill, "Adults and video games," Pew Internet and American Life Project, 2008, Available: www.pewinternet.org.

[4] Entertainment Software Association (ESA). Essential facts about the computer and video game industry. ESA EF 2013.

[5] W. R. Boot, A. F. Kramer, D. J. Simons, M. Fabiani, G. Gratton, "The effects of video game playing on attention, memory, and executive control," Acta Psychologica, 2008, vol. 129, pp. 387–398.

[6] J.E. Clark, A. K. Lanphear, C. C. Riddick, "The Effects of Videogame Playing on the Response Selection processing of Elderly Adults," Journal of Gerontology, 1987, vol. 42, pp. 82–85.

[7] K. Rayner, "Eye Movements in Reading and Information Processing: 20 Years of Research," Psychological Bulletin, 1998, vol. 124(3), pp. 372–422.

[8] A.Yarbus, "Eye movements and vision," Plenum Press, 1967.

[9] W. J. Gibbs, R. S. Bernas, "Visual Attention in Newspaper versus TV-Oriented News Websites," Journal of Usability Studies, 2009, vol.4, pp.147–165.

[10] J. H. Goldberg and M. J. Stimson and M. Lewenstein and N. Scottand and A. M Wichansky, "Eye tracking in web search tasks: design implications," Proc. of the 2002 symposium onEye tracking research & applications (ETRA 02), 2002, pp. 51–58, doi:10.1145/507072.507082.

[11] S. Lee, G. J. Kim, S. Choi, "Real-time tracking of visually attended objects in interactive virtual environments," Proc. ACM Symp. Virtual reality software and technology (VRST 07), 2007, ACM, pp. 29–38, doi:10.1145/1315184.1315187.

[12] Y. Kenji, W. Katsumi, K. Takashi, "Dynamic evaluation of distribution of visual attention during playing video game," Proc. of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology (ACE 06), 2006, Article No. 96, doi:10.1145/1178823.1178934.

[13] P. Cubaud, A. Topol, J. Dupire, "Using Game Engines for non 3D Gaming Applications," Proc. of the international conference on computer games (CGAMES 05), 2005, pp. 304–307.

[14] M. S. El-Nasr, S. Yan, "Visual attention in 3D video games," Proc. of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology (ACE 06), ACM, 2006, article No. 22, doi:101145/1178823.1178849.

[15] A. Kenny, H. Koesling , D. Delaney , S. Mcloone , T. Ward, "A Preliminary Investigation into Eye Gaze Data in a First Person Shooter Game," Proc. 19th European Conference Modelling and Simulation (ECMS), 2005,vol. 5, pp. 146–152.

[16] P. M. Fischer, J. W. Richards, E. J. Berman, D. M. Krugman, "Recall and Eye Tracking Study of Adolescents Viewing Tobacco Advertisements," Journal of the American Medical Association, 1989, vol. 261 (1), pp. 84–89, doi: 10.1001/jama.261.1.84.

[17] L. J.Muir, I. Richardson, "Perception of sign language and its application to visual communications for deaf people," The Journal of Deaf Studies and Deaf Education, vol. 10(4), 2005, pp. 390–401, doi:10.1093/deafed/eni037.

[18] R. J. K. Jacob, K. S. Karn, "Eye tracking in Human-Computer Interaction and usability research: Ready to deliver the promises," The mind's eye: Cognitive and applied aspects of eye movement research, Elsevier, 2003, pp. 573–605.

[19] S. Hillaire , A. Lecuyer, T. Regia-Corte , R. Cozot , J. Royan, G. Breton, "A real-time visual attention model for predicting gaze point during first-person exploration of virtual environments," Proc. ACM Symp. Virtual Reality Software and Technology (VRST 10), ACM, 2010, pp. 191–198, doi:10.1145/1889863.1889907.

[20] K. K. Ball, B. L. Beard, D. L. Roenker, R. L. Miller, D. S. Griggs, "Age and visual search: Expanding the useful field of view," Journal of the Optical Society of America, 1988, vol. 5, pp. 2210–2220, doi:10.1364/JOSAA.5.002210.

[21] K. Murphy, A. Spencer, "Playing video games does not make for better visual attention skills," Journal of Articles in Support of the Null Hypothesis, 2009, JASNH, vol. 6, No. 1.